

Erasure Coding for Real-Time Streaming

Derek Leong

Department of Electrical Engineering
California Institute of Technology
Pasadena, California 91125, USA
derekleong@caltech.edu

Tracey Ho

Department of Electrical Engineering
California Institute of Technology
Pasadena, California 91125, USA
tho@caltech.edu

Abstract—We consider a real-time streaming system where messages are created sequentially at the source, and are encoded for transmission to the receiver over a packet erasure link. Each message must subsequently be decoded at the receiver within a given delay from its creation time. The goal is to construct an erasure correction code that achieves the maximum message size when all messages must be decoded by their respective deadlines under a specified set of erasure patterns (erasure model). We present an explicit intrasession code construction that is asymptotically optimal under erasure models containing a limited number of erasures per coding window, per sliding window, and containing erasure bursts of a limited length.

I. INTRODUCTION

We consider a real-time streaming system where messages are created sequentially at the source, and are encoded for transmission to the receiver over a packet erasure link. Each message must subsequently be decoded at the receiver within a given delay from its creation time. We seek to construct an erasure correction code that withstands a specified set of erasure patterns (erasure model), allowing all messages to be decoded by their respective deadlines.

In particular, we consider three erasure models: the first model limits the number of erasures in each coding window, the second model limits the number of erasures in each sliding window, while the third model limits the length of erasure bursts. For each erasure model, the objective is to find an optimal code that achieves the maximum message size, among all codes that allow all messages to be decoded by their respective deadlines under all admissible erasure patterns.

We present an explicit intrasession code construction which specifies an allocation of link bandwidth or data packet space among the different messages; coding occurs within each message but not across messages. Intrasession coding is attractive due to its relative simplicity, but it is not known in general when intrasession coding is sufficient or when intersession coding is necessary. We show that for an asymptotic number of messages, our code construction achieves the optimal message size among all codes (intrasession or intersession) for the first and second models with any given maximum number of

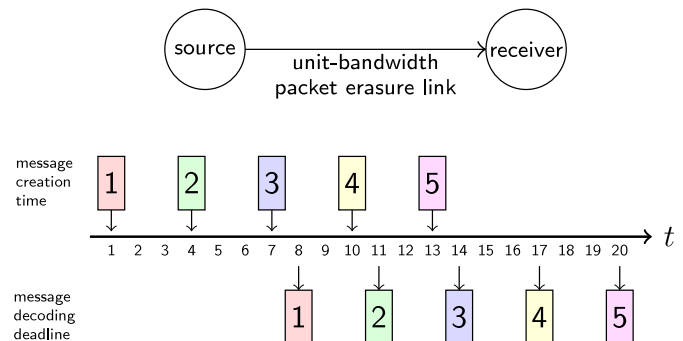


Fig. 1. Real-time streaming system for $(c, d) = (3, 8)$. Each of the messages $\{1, \dots, 5\}$ is assigned a unique color. Messages are created at the source at regular intervals of c time steps, and must be decoded at the receiver within d time steps from their respective creation times. At each time step t , the source is allowed to transmit a single data packet of normalized unit size over the link.

erasures per window, and for the third model when the given maximum erasure burst length is sufficiently short or long.

In related work, Martinian *et al.* [2], [3] provided constructions of streaming codes that minimize the delay required to correct erasure bursts of a given length. Streaming codes in which the decoding error probability decays exponentially with delay, called tree codes or anytime codes, are considered in [4]–[6]. Tekin *et al.* [7] considered erasure correction coding for a non-real-time streaming system where all messages are initially present at the encoder.

We begin with a formal definition of the problem in Section II, and proceed to describe the construction of our intrasession code in Section III. We then demonstrate the optimality of this code under each erasure model in the subsequent sections. Proofs of theorems are deferred to Appendix B.

II. PROBLEM DEFINITION

Consider a discrete-time data streaming system comprising a source and a receiver, with a directed unit-bandwidth packet erasure link from the source to the receiver. Independent messages of uniform size $s > 0$ are created at the source at regular intervals of $c \in \mathbb{Z}^+$ time steps, and must be decoded at the receiver within $d \in \mathbb{Z}^+$ time steps from their respective creation times. At each time step $t \in \mathbb{Z}^+$, the source is allowed to transmit a single data packet of normalized unit size over the link. Fig. 1 depicts this real-time streaming system for an

This paper is an extended version of [1], which was presented at the ISIT 2012 conference.

This work was supported in part by the Air Force Office of Scientific Research under Grant FA9550-10-1-0166.

instance of (c, d) .

More precisely, each message $k \in \mathbb{Z}^+$ is created at time step $(k-1)c+1$, and is to be decoded by time step $(k-1)c+d$. The coded data transmitted at each time step $t \in \mathbb{Z}^+$ must be a function of messages created at time step t or earlier. Let coding window W_k be the interval of d time steps between the creation time and the decoding deadline of message k , i.e.,

$$W_k \triangleq \{(k-1)c+1, \dots, (k-1)c+d\}.$$

We shall assume that $d > c$ so as to avoid the degenerate case of nonoverlapping coding windows for which it is sufficient to code individual messages separately.

Consider the first n messages $\{1, \dots, n\}$, and the union of their (overlapping) coding windows T_n given by

$$T_n \triangleq W_1 \cup \dots \cup W_n = \{1, \dots, (n-1)c+d\}.$$

An erasure pattern $E \subseteq T_n$ specifies a set of erased data transmissions over the link. More precisely, if $t \in E$, then none of the data transmitted at time step t is received by the receiver (i.e., the entire data packet is erased); if $t \in T_n \setminus E$, then all of the data transmitted at time step t is received by the receiver at time step t (i.e., the entire data packet is received without delay). An erasure model specifies a set of erasure patterns that an erasure correction code should withstand.

For a given pair of positive integers a and b , we define the offset quotient $q_{a,b}$ and remainder $r_{a,b}$ to be the unique integers satisfying the following three conditions:

$$a = q_{a,b} b + r_{a,b}, \quad q_{a,b} \in \mathbb{Z}_0^+, \quad r_{a,b} \in \{1, \dots, b\},$$

where \mathbb{Z}_0^+ denotes the set of nonnegative integers, i.e., $\mathbb{Z}^+ \cup \{0\}$. Note that our definition departs from the usual definition of quotient and remainder in that $r_{a,b}$ can be equal to b but not 0.

III. CODE CONSTRUCTION

We present an intrasession code which codes only within each message and not across different messages. We begin by specifying the amount of link bandwidth or data packet space allocated for the encoding of each message at each time step. An appropriate code (e.g., random linear coding, MDS code) is then applied to the allocation so that each message can be decoded whenever the total amount of received data that encodes that message is at least the message size s .

The allocation of link bandwidth follows a simple rule: the link bandwidth at each time step is divided evenly among all *active* messages. We say that message k is active at time step t if and only if t falls within its coding window, i.e., $t \in W_k$. Fig. 2 shows how much link bandwidth at each time step is allocated to each message, for two instances of (c, d) .

For a given choice of (c, d) , the messages that are encoded at a given time step $t \in \mathbb{Z}^+$ can be stated explicitly as follows: First, we define A_t to be the set of active messages at time step t , i.e.,

$$\begin{aligned} A_t &\triangleq \{k \in \mathbb{Z}^+ : t \in W_k\} \\ &= \{k \in \mathbb{Z}^+ : (k-1)c+1 \leq t \leq (k-1)c+d\} \end{aligned}$$

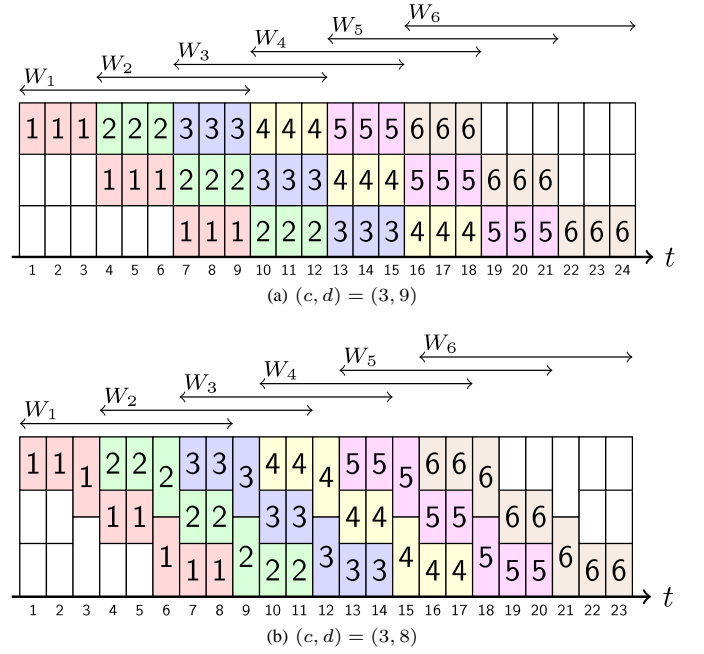


Fig. 2. Allocation of link bandwidth at each time step t , in the encoding of messages $\{1, \dots, 6\}$, for (a) $(c, d) = (3, 9)$ and (b) $(c, d) = (3, 8)$. Each message is assigned a unique color. In (a), because d is a multiple of c , we have $q_{d,c} + 1 = 3$ active messages at each time step. In (b), because d is not a multiple of c , we have either $q_{d,c} = 2$ or $q_{d,c} + 1 = 3$ active messages at each time step.

$$= \left\{ k \in \mathbb{Z}^+ : \frac{t-d}{c} + 1 \leq k \leq \frac{t-1}{c} + 1 \right\}.$$

Treating nonpositive messages $0, -1, -2, \dots$ as dummy messages, we can write

$$A_t = \left\{ \left\lceil \frac{t-d}{c} + 1 \right\rceil, \dots, \left\lfloor \frac{t-1}{c} + 1 \right\rfloor \right\}.$$

Expressing this in terms of $q_{d,c}$, $r_{d,c}$, $q_{t,c}$, $r_{t,c}$ yields

$$A_t = \left\{ q_{t,c} + 1 - q_{d,c} + \left\lceil \frac{r_{t,c} - r_{d,c}}{c} \right\rceil, \dots, q_{t,c} + 1 \right\}.$$

It follows that the number of active messages $|A_t|$ varies over time depending on the value of $r_{t,c}$; specifically, two cases are possible:

Case 1: If $r_{t,c} \leq r_{d,c}$, then

$$-1 < \frac{1-c}{c} \leq \frac{r_{t,c} - r_{d,c}}{c} \leq 0,$$

which implies that $\left\lceil \frac{r_{t,c} - r_{d,c}}{c} \right\rceil = 0$, and

$$A_t = \{q_{t,c} + 1 - q_{d,c}, \dots, q_{t,c} + 1\}.$$

The $q_{d,c} + 1$ messages of A_t are therefore encoded at time step t , with each message allocated $\frac{1}{q_{d,c}+1}$ amount of link bandwidth.

Case 2: If $r_{t,c} > r_{d,c}$, then

$$0 < \frac{r_{t,c} - r_{d,c}}{c} \leq \frac{c-1}{c} < 1,$$

which implies that $\left\lceil \frac{r_{t,c} - r_{d,c}}{c} \right\rceil = 1$, and

$$A_t = \{q_{t,c} + 1 - (q_{d,c} - 1), \dots, q_{t,c} + 1\}.$$

The $q_{d,c}$ messages of A_t are therefore encoded at time step t , with each message allocated $\frac{1}{q_{d,c}}$ amount of link bandwidth.

Note that when d is a multiple of c , we have $r_{t,c} \leq r_{d,c} = c$ for any t , which implies that $q_{d,c} + 1$ messages are encoded at every time step.

In our subsequent performance analysis of this code, we make repeated use of two key code properties; these are presented as technical lemmas in Appendix A.

IV. PERFORMANCE UNDER z ERASURES PER CODING WINDOW

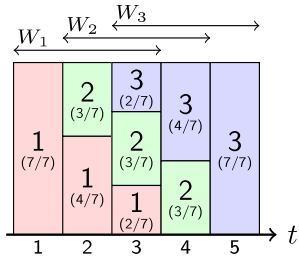
For the first erasure model, we look at erasure patterns that have a limited number of erasures per coding window. Consider the first n messages $\{1, \dots, n\}$, and the union of their (overlapping) coding windows T_n . Let $\mathcal{E}_n^{\text{CW}}$ be the set of erasure patterns that have z or fewer erasures in each coding window W_k , i.e.,

$$\mathcal{E}_n^{\text{CW}} \triangleq \{E \subseteq T_n : |E \cap W_k| \leq z \forall k \in \{1, \dots, n\}\}.$$

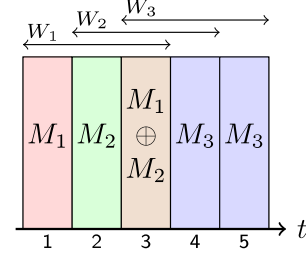
The objective is to construct a code that allows all n messages $\{1, \dots, n\}$ to be decoded by their respective deadlines under any erasure pattern $E \in \mathcal{E}_n^{\text{CW}}$. Let s_n^{CW} be the maximum message size that can be achieved by such a code, for a given choice of (n, c, d, z) .

We observe that over a finite time horizon (i.e., when the number of messages n is finite), intrasession coding can be suboptimal. The following example shows that an intersession code can achieve a message size that is strictly larger than the message size achieved by an optimal intrasession code:

Example (Finite time horizon). Suppose that $(n, c, d, z) = (3, 1, 3, 1)$. The maximum message size that can be achieved by an intrasession code is $\frac{6}{7}$; one such optimal intrasession code, which can be found by solving a linear program, is as follows (the amount of link bandwidth allocated to each message is indicated in parentheses):



The following intersession code achieves a strictly larger message size of 1 (M_k denotes message k):



Using a simple cut-set bound argument, we can show that this is also the maximum achievable message size, i.e., $s_n^{\text{CW}} = 1$.

However, it turns out that the intrasession code constructed in Section III is *asymptotically* optimal; the gap between the maximum achievable message size s_n^{CW} and the message size achieved by the code vanishes as the number of messages n goes to infinity:

Theorem 1. *The code constructed in Section III is asymptotically optimal in the following sense: the code achieves a message size of*

$$\sum_{j=1}^{d-z} y_j,$$

which is equal to the asymptotic maximum achievable message size $\lim_{n \rightarrow \infty} s_n^{\text{CW}}$, where $\mathbf{y} = (y_1, \dots, y_d)$ is defined as

$$\mathbf{y} \triangleq \left(\underbrace{\frac{1}{q_{d,c}+1}, \dots, \frac{1}{q_{d,c}+1}}_{(q_{d,c}+1)r_{d,c} \text{ entries}}, \underbrace{\frac{1}{q_{d,c}}, \dots, \frac{1}{q_{d,c}}}_{q_{d,c}(c-r_{d,c}) \text{ entries}} \right).$$

The achievability claim of this theorem is a consequence of Lemma 1; to prove the converse claim, we consider a cut-set bound corresponding to a specific *worst-case* erasure pattern in which exactly z erasures occur in every coding window. This erasure pattern is chosen with the help of Lemma 2; specifically, the erased time steps are chosen to coincide with the larger blocks allocated to each message in the constructed code.

V. PERFORMANCE UNDER z ERASURES PER SLIDING WINDOW OF d TIME STEPS

For the second erasure model, we look at erasure patterns that have a limited number of erasures per sliding window of d time steps. Consider the first n messages $\{1, \dots, n\}$, and the union of their (overlapping) coding windows T_n . Let sliding window L_t denote the interval of d time steps beginning at time step t , i.e.,

$$L_t \triangleq \{t, \dots, t + d - 1\}.$$

Let $\mathcal{E}_n^{\text{SW}}$ be the set of erasure patterns that have z or fewer erasures in each sliding window L_t , i.e.,

$$\mathcal{E}_n^{\text{SW}} \triangleq \{E \subseteq T_n : |E \cap L_t| \leq z \forall t \in \{1, \dots, (n-1)c + 1\}\}.$$

The objective is to construct a code that allows all n messages $\{1, \dots, n\}$ to be decoded by their respective deadlines under any erasure pattern $E \in \mathcal{E}_n^{\text{SW}}$. Let s_n^{SW} be the maximum message

size that can be achieved by such a code, for a given choice of (n, c, d, z) .

We note that since $\mathcal{E}_n^{\text{SW}} \subseteq \mathcal{E}_n^{\text{CW}}$, we therefore have $s_n^{\text{SW}} \geq s_n^{\text{CW}}$. For the special case of $c = 1$, each sliding window is also a coding window, and so this sliding window erasure model reduces to the coding window erasure model of Section IV, i.e., $\mathcal{E}_n^{\text{SW}} = \mathcal{E}_n^{\text{CW}}$. Over a finite time horizon, intrasession coding can also be suboptimal for this erasure model; the illustrating example from Section IV applies here as well.

Surprisingly, the constructed intrasession code also turns out to be asymptotically optimal over all codes; the omission of erasure patterns in $\mathcal{E}_n^{\text{SW}}$ compared to $\mathcal{E}_n^{\text{CW}}$ has not led to an increase in the maximum achievable message size (cf. Theorem 1):

Theorem 2. *The code constructed in Section III is asymptotically optimal in the following sense: the code achieves a message size of*

$$\sum_{j=1}^{d-z} y_j,$$

which is equal to the asymptotic maximum achievable message size $\lim_{n \rightarrow \infty} s_n^{\text{SW}}$.

Proving the converse claim of this theorem requires a different approach from that of Theorem 1. When d is a multiple of c , we need only consider a cut-set bound corresponding to an obvious *worst-case* erasure pattern in which exactly z erasures occur in every sliding window, specifically, a periodic erasure pattern with alternating intervals of z erased time steps and $d - z$ unerased time steps. When d is not a multiple of c , no single admissible erasure pattern provides a cut-set bound that matches the constructed code; instead, we need to combine different erasure patterns for different messages. To pick these erasure patterns, we first choose a specific *base* erasure pattern E' (which may not be admissible in general) with the help of Lemma 2. We then derive admissible erasure patterns from E' by taking its intersection with each coding window, i.e., $(E' \cap W_k) \in \mathcal{E}_n^{\text{SW}}$. These derived erasure patterns are used in the inductive computation of an upper bound for the conditional entropy

$$H(X[W_n \setminus E'] \mid M_1^n, X_1^{(n-1)c}),$$

where X_t is a random variable representing the coded data transmitted at time step t , M_k is a random variable representing message k , and $X[A] \triangleq (X_t)_{t \in A}$. Intuitively, this conditional entropy term expresses how much space is left in the unerased data packets of the coding window for message n , after encoding the first n messages, and conditioned on the previous time steps. The nonnegativity of the conditional entropy leads us to a bound for s_n^{SW} that matches the message size achieved by the constructed code in the limit $n \rightarrow \infty$.

VI. PERFORMANCE UNDER ERASURE BURSTS OF z TIME STEPS

For the third erasure model, we look at erasure patterns that contain erasure bursts of a limited number of time steps.

Consider the first n messages $\{1, \dots, n\}$, and the union of their (overlapping) coding windows T_n . Let \mathcal{E}_n^{B} be the set of erasure patterns in which each erasure burst is z or fewer time steps in length, and consecutive bursts are separated by a gap of $d - z$ or more unerased time steps, i.e.,

$$\mathcal{E}_n^{\text{B}} \triangleq \left\{ E \subseteq T_n : \begin{aligned} (t \in E \wedge t+1 \notin E) &\Rightarrow |E \cap \{t+1, \dots, t+d-z\}| = 0, \\ (t \notin E \wedge t+1 \in E) &\Rightarrow |E \cap \{t+1, \dots, t+z+1\}| \leq z \end{aligned} \right\}.$$

The objective is to construct a code that allows all n messages $\{1, \dots, n\}$ to be decoded by their respective deadlines under any erasure pattern $E \in \mathcal{E}_n^{\text{B}}$. Let s_n^{B} be the maximum message size that can be achieved by such a code, for a given choice of (n, c, d, z) .

Using the proof technique of Theorem 2, we can show that the constructed intrasession code is asymptotically optimal when d is a multiple of c , or when the maximum erasure burst length z is sufficiently short or long:

Theorem 3. *If*

- 1) d is a multiple of c , or
- 2) d is not a multiple of c and $z \leq c - r_{d,c}$, or
- 3) d is not a multiple of c and $z \geq d - r_{d,c} = q_{d,c}c$,

then the code constructed in Section III is asymptotically optimal in the following sense: the code achieves a message size of

$$\sum_{j=1}^{d-z} y_j,$$

which is equal to the asymptotic maximum achievable message size $\lim_{n \rightarrow \infty} s_n^{\text{B}}$.

When the maximum erasure burst length z takes on intermediate values, intersession coding may become necessary. We are currently studying optimal convolutional codes for this case.

APPENDIX A CODE PROPERTIES

The first property describes when it is possible to decode each message:

Lemma 1 (Achievability). *Consider the code constructed in Section III for a given choice of (c, d) . If message size s satisfies the inequality*

$$s \leq \sum_{j=1}^{\ell} y_j,$$

where $\mathbf{y} = (y_1, \dots, y_d)$ is as defined in Theorem 1, then each message $k \in \mathbb{Z}^+$ can be decoded from the data at any ℓ time steps in its coding window W_k .

Note that the maximum message size s that can be supported by the code is given by $\sum_{j=1}^d y_j = c$, which corresponds to the choice of $\ell = d$.

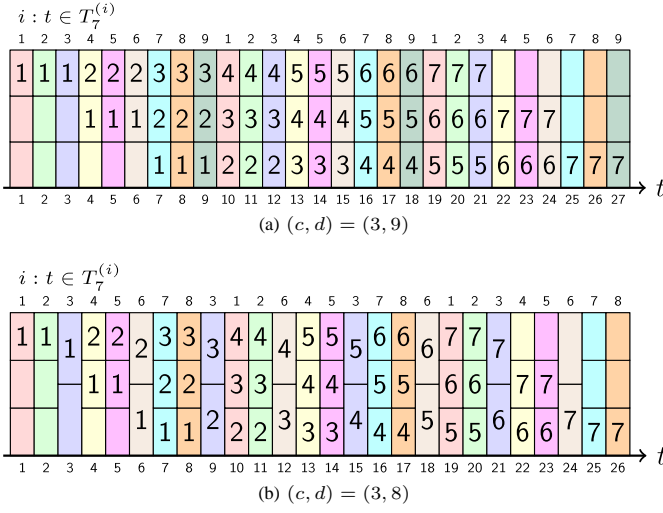


Fig. 3. Partitioning of the set of time steps T_n into the d sets $T_n^{(1)}, \dots, T_n^{(d)}$, in the encoding of messages $\{1, \dots, 7\}$, for (a) $(c, d) = (3, 9)$ and (b) $(c, d) = (3, 8)$. Each set $T_n^{(i)}$ is assigned a unique color. The number i at the top of each time step t indicates the set $T_n^{(i)}$ to which t belongs.

The second property describes a way of partitioning time steps into sets with certain specific properties, which are used in our specification of the worst-case erasure patterns:

Lemma 2 (Partition of Coding Windows). *Consider the code constructed in Section III for a given choice of (c, d) . Consider the first n messages $\{1, \dots, n\}$, and the union of their (overlapping) coding windows T_n . The set of time steps T_n can be partitioned into d sets $T_n^{(1)}, \dots, T_n^{(d)}$, given by*

$$T_n^{(i)} \triangleq \begin{cases} \left\{ (j(q_{d,c} + 1) + q_{i,c})c + r_{i,c} \in T_n : j \in \mathbb{Z}_0^+ \right\} & \text{if } r_{i,c} \leq r_{d,c}, \\ \left\{ (j q_{d,c} + q_{i,c})c + r_{i,c} \in T_n : j \in \mathbb{Z}_0^+ \right\} & \text{if } r_{i,c} > r_{d,c}, \end{cases}$$

with the following properties:

- 1) Over the time steps in the set $T_n^{(i)}$, each message $k \in \{1, \dots, n\}$ is allocated $\frac{1}{q_{d,c}+1}$ amount of link bandwidth if $r_{i,c} \leq r_{d,c}$, and $\frac{1}{q_{d,c}}$ amount of link bandwidth if $r_{i,c} > r_{d,c}$.
- 2) The allocated link bandwidth in $T_n^{(i)}$ for each message $k \in \{1, \dots, n\}$ is contained within a single time step in its coding window W_k (as opposed to being spread over multiple time steps or being outside of the coding window).
- 3) The total amount of link bandwidth over all time steps in $T_n^{(i)}$, i.e., $|T_n^{(i)}|$, has the following upper bound:

$$|T_n^{(i)}| < \begin{cases} \frac{n}{q_{d,c} + 1} + 2 & \text{if } r_{i,c} \leq r_{d,c}, \\ \frac{n}{q_{d,c}} + 2 & \text{if } r_{i,c} > r_{d,c}. \end{cases}$$

Fig. 3 shows how the set of time steps T_n is partitioned into the d sets $T_n^{(1)}, \dots, T_n^{(d)}$, for two instances of (c, d) .

APPENDIX B PROOFS OF THEOREMS

Proof of Lemma 1: Consider a given message $k \in \mathbb{Z}^+$ and its coding window

$$W_k = \{(k-1)c + i : i \in \{1, \dots, d\}\}.$$

At each time step $t \in W_k$, message k is allocated either $\frac{1}{q_{d,c}+1}$ or $\frac{1}{q_{d,c}}$ amount of link bandwidth; at all other time steps $t \notin W_k$, message k is allocated zero link bandwidth.

Let x_i be the amount of link bandwidth at time step $t = (k-1)c + i$ that is allocated to message k . Writing t in terms of $q_{i,c}$ and $r_{i,c}$ produces

$$t = (k-1)c + i = \underbrace{(k-1 + q_{i,c})c}_{q_{t,c}} + \underbrace{r_{i,c}}_{r_{t,c}}.$$

It follows from the code construction that the value of x_i depends on $r_{i,c}$; specifically, two cases are possible:

Case 1: If $r_{i,c} \leq r_{d,c}$, then $x_i = \frac{1}{q_{d,c}+1}$. Since $i \in \{1, \dots, d\}$, this condition corresponds to the case where $q_{i,c} \in \{0, \dots, q_{d,c}\}$ and $r_{i,c} \in \{1, \dots, r_{d,c}\}$. Therefore, message k is allocated $\frac{1}{q_{d,c}+1}$ amount of link bandwidth per time step for a total of $(q_{d,c} + 1)r_{d,c}$ time steps in the coding window W_k .

Case 2: If $r_{i,c} > r_{d,c}$, then $x_i = \frac{1}{q_{d,c}}$. Since $i \in \{1, \dots, d\}$, this condition corresponds to the case where $q_{i,c} \in \{0, \dots, q_{d,c} - 1\}$ and $r_{i,c} \in \{r_{d,c} + 1, \dots, c\}$. Therefore, message k is allocated $\frac{1}{q_{d,c}}$ amount of link bandwidth per time step for a total of $q_{d,c}(c - r_{d,c})$ time steps in the coding window W_k .

Observe that \mathbf{y} is simply a vector containing the elements of $\{x_i\}_{i=1}^d$ sorted in ascending order. Since

$$\sum_{i \in U} x_i \geq \sum_{j=1}^{|U|} y_j \quad \forall U \subseteq \{1, \dots, d\},$$

it follows that over any ℓ time steps in the coding window W_k , the total amount of link bandwidth allocated to message k is at least $\sum_{j=1}^{\ell} y_j$. Therefore, as long as the message size s does not exceed $\sum_{j=1}^{\ell} y_j$, message k can always be decoded from the data at any ℓ time steps in W_k . ■

Proof of Lemma 2: The stated partition can be constructed by assigning each time step $t \in T_n$ to the set $T_n^{(q_{i,c}c + r_{i,c})}$, where

$$r_{i,c} = r_{t,c},$$

$$q_{i,c} = \begin{cases} q_{t,c} - \left\lfloor \frac{q_{t,c}}{q_{d,c}+1} \right\rfloor (q_{d,c} + 1) & \text{if } r_{t,c} \leq r_{d,c}, \\ q_{t,c} - \left\lfloor \frac{q_{t,c}}{q_{d,c}} \right\rfloor q_{d,c} & \text{if } r_{t,c} > r_{d,c}. \end{cases}$$

Note that index $q_{i,c}c + r_{i,c} \in \{1, \dots, d\}$ since $q_{i,c} \in \{0, \dots, q_{d,c}\}$ when $r_{i,c} \in \{1, \dots, r_{d,c}\}$, and $q_{i,c} \in \{0, \dots, q_{d,c} - 1\}$ when $r_{i,c} \in \{r_{d,c} + 1, \dots, c\}$. To prove the required code properties, we consider two separate cases:

Case 1: Consider the set $T_n^{(i)}$ for a choice of i satisfying $r_{i,c} \leq r_{d,c}$. Since each time step $t \in T_n^{(i)}$ can be expressed as

$$t = \underbrace{(j(q_{d,c} + 1) + q_{i,c})}_q c + \underbrace{r_{i,c}}_{r_{t,c}} \triangleq t_j, \quad \text{where } j \in \mathbb{Z}_0^+,$$

it follows from the code construction that the set of active messages at each time step contains $q_{d,c} + 1$ messages, and is given by

$$A_{t_j} = \left\{ \underbrace{j(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c}}_{q_{t,c}}, \dots, \underbrace{j(q_{d,c} + 1) + q_{i,c} + 1}_{q_{t,c}} \right\}.$$

The smallest time step in $T_n^{(i)}$ corresponds to the choice of $j = 0$, which produces $t_0 = q_{i,c}c + r_{i,c} = i$ and the set of active messages

$$A_{t_0} = \{q_{i,c} + 1 - q_{d,c}, \dots, q_{i,c} + 1\}.$$

Note that A_{t_0} contains message 1 since $q_{i,c} \in \{0, \dots, q_{d,c}\}$, which implies that

$$q_{i,c} + 1 - q_{d,c} \leq 1 \leq q_{i,c} + 1.$$

At the other extreme, let the largest time step in $T_n^{(i)}$ correspond to the choice of $j = j'$; we therefore have

$$t_{j'} \leq (n - 1)c + d < t_{j'+1}, \quad (1)$$

and the final set of active messages

$$A_{t_{j'}} = \{j'(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c}, \dots, j'(q_{d,c} + 1) + q_{i,c} + 1\}.$$

From the first inequality of (1), we obtain

$$\begin{aligned} (j'(q_{d,c} + 1) + q_{i,c})c + r_{i,c} &\leq (n - 1 + q_{d,c})c + r_{d,c} \\ \Rightarrow n &\geq \left\lceil \frac{(j'(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c})c + r_{i,c} - r_{d,c}}{c} \right\rceil \\ &= j'(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c} + \left\lceil \frac{r_{i,c} - r_{d,c}}{c} \right\rceil \\ &= j'(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c}, \end{aligned} \quad (2)$$

where the final step follows from the fact that $1 \leq r_{i,c} \leq r_{d,c} \leq c$, which implies that

$$-1 < \frac{1 - c}{c} \leq \frac{r_{i,c} - r_{d,c}}{c} \leq 0 \quad \Rightarrow \quad \left\lceil \frac{r_{i,c} - r_{d,c}}{c} \right\rceil = 0.$$

From the second inequality of (1), we obtain

$$\begin{aligned} (n - 1 + q_{d,c})c + r_{d,c} &\leq ((j' + 1)(q_{d,c} + 1) + q_{i,c})c + r_{i,c} - 1 \\ \Rightarrow n &\leq \left\lfloor \frac{((j' + 1)(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c})c + r_{i,c} - r_{d,c} - 1}{c} \right\rfloor \\ &= (j' + 1)(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c} + \left\lfloor \frac{r_{i,c} - r_{d,c} - 1}{c} \right\rfloor \\ &= j'(q_{d,c} + 1) + q_{i,c} + 1, \end{aligned} \quad (3)$$

where the final step follows from the fact that $1 \leq r_{i,c} \leq r_{d,c} \leq c$, which implies that

$$-1 = \frac{1 - c - 1}{c} \leq \frac{r_{i,c} - r_{d,c} - 1}{c} \leq -\frac{1}{c} < 0$$

$$\Rightarrow \left\lfloor \frac{r_{i,c} - r_{d,c} - 1}{c} \right\rfloor = -1.$$

By combining inequalities (2) and (3), we arrive at

$$j'(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c} \leq n \leq j'(q_{d,c} + 1) + q_{i,c} + 1,$$

which enables us to infer that $A_{t_{j'}}$ contains message n .

For any pair of consecutive time steps $t_j, t_{j+1} \in T_n^{(i)}$, where

$$\begin{aligned} t_j &= (j(q_{d,c} + 1) + q_{i,c})c + r_{i,c}, \\ t_{j+1} &= ((j + 1)(q_{d,c} + 1) + q_{i,c})c + r_{i,c}, \end{aligned}$$

we observe that the smallest message in $A_{t_{j+1}}$ is exactly one larger than the largest message in A_{t_j} , i.e.,

$$\begin{aligned} (j + 1)(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c} \\ &= j(q_{d,c} + 1) + q_{i,c} + 1 - q_{d,c} + q_{d,c} + 1 \\ &= (j(q_{d,c} + 1) + q_{i,c} + 1) + 1. \end{aligned}$$

Thus, there are no overlapping or omitted messages among the sets of active messages corresponding to $T_n^{(i)}$. Properties 1 and 2 therefore follow.

The total amount of link bandwidth over all time steps in $T_n^{(i)}$, i.e., $|T_n^{(i)}|$, can be computed by summing over the link bandwidth allocated to the n messages, and adding the unused link bandwidth in the smallest time step (which is allocated to nonpositive dummy messages) and in the largest time step (which is allocated to messages larger than n); this produces the required upper bound of Property 3.

Case 2: Consider the set $T_n^{(i)}$ for a choice of i satisfying $r_{i,c} > r_{d,c}$. Since each time step $t \in T_n^{(i)}$ can be expressed as

$$t = \underbrace{(j q_{d,c} + q_{i,c})}_q c + \underbrace{r_{i,c}}_{r_{t,c}} \triangleq t_j, \quad \text{where } j \in \mathbb{Z}_0^+,$$

it follows from the code construction that the set of active messages at each time step contains $q_{d,c}$ messages, and is given by

$$A_{t_j} = \left\{ \underbrace{j q_{d,c} + q_{i,c} + 1 - (q_{d,c} - 1)}_{q_{t,c}}, \dots, \underbrace{j q_{d,c} + q_{i,c} + 1}_{q_{t,c}} \right\}.$$

The smallest time step in $T_n^{(i)}$ corresponds to the choice of $j = 0$, which produces $t_0 = q_{i,c}c + r_{i,c} = i$ and the set of active messages

$$A_{t_0} = \{q_{i,c} + 1 - (q_{d,c} - 1), \dots, q_{i,c} + 1\}.$$

Note that A_{t_0} contains message 1 since $q_{i,c} \in \{0, \dots, q_{d,c} - 1\}$, and therefore

$$q_{i,c} + 1 - (q_{d,c} - 1) \leq 1 \leq q_{i,c} + 1.$$

At the other extreme, let the largest time step in $T_n^{(i)}$ correspond to the choice of $j = j'$; we therefore have

$$t_{j'} \leq (n - 1)c + d < t_{j'+1}, \quad (4)$$

and the final set of active messages

$$A_{t_{j'}} = \{j' q_{d,c} + q_{i,c} + 1 - (q_{d,c} - 1), \dots, j' q_{d,c} + q_{i,c} + 1\}.$$

From the first inequality of (4), we obtain

$$\begin{aligned}
(j' q_{d,c} + q_{i,c})c + r_{i,c} &\leq (n - 1 + q_{d,c})c + r_{d,c} \\
\Rightarrow n &\geq \left\lceil \frac{(j' q_{d,c} + q_{i,c} + 1 - q_{d,c})c + r_{i,c} - r_{d,c}}{c} \right\rceil \\
&= j' q_{d,c} + q_{i,c} + 1 - q_{d,c} + \left\lceil \frac{r_{i,c} - r_{d,c}}{c} \right\rceil \\
&= j' q_{d,c} + q_{i,c} + 1 - (q_{d,c} - 1), \tag{5}
\end{aligned}$$

where the final step follows from the fact that $1 \leq r_{d,c} < r_{i,c} \leq c$, which implies that

$$0 < \frac{r_{i,c} - r_{d,c}}{c} \leq \frac{c-1}{c} < 1 \Rightarrow \left\lceil \frac{r_{i,c} - r_{d,c}}{c} \right\rceil = 1.$$

From the second inequality of (4), we obtain

$$\begin{aligned}
(n - 1 + q_{d,c})c + r_{d,c} &\leq ((j' + 1)q_{d,c} + q_{i,c})c + r_{i,c} - 1 \\
\Rightarrow n &\leq \left\lfloor \frac{((j' + 1)q_{d,c} + q_{i,c} + 1 - q_{d,c})c + r_{i,c} - r_{d,c} - 1}{c} \right\rfloor \\
&= (j' + 1)q_{d,c} + q_{i,c} + 1 - q_{d,c} + \left\lfloor \frac{r_{i,c} - r_{d,c} - 1}{c} \right\rfloor \\
&= j' q_{d,c} + q_{i,c} + 1, \tag{6}
\end{aligned}$$

where the final step follows from the fact that $1 \leq r_{d,c} < r_{i,c} \leq c$, which implies that

$$\begin{aligned}
0 = \frac{1-1}{c} &\leq \frac{r_{i,c} - r_{d,c} - 1}{c} \leq \frac{c-1-1}{c} < 1 \\
\Rightarrow \left\lfloor \frac{r_{i,c} - r_{d,c} - 1}{c} \right\rfloor &= 0.
\end{aligned}$$

By combining inequalities (5) and (6), we arrive at

$$j' q_{d,c} + q_{i,c} + 1 - (q_{d,c} - 1) \leq n \leq j' q_{d,c} + q_{i,c} + 1,$$

which enables us to infer that $A_{t_{j'}}$ contains message n .

For any pair of consecutive time steps $t_j, t_{j+1} \in T_n^{(i)}$, where

$$\begin{aligned}
t_j &= (j q_{d,c} + q_{i,c})c + r_{i,c}, \\
t_{j+1} &= ((j + 1) q_{d,c} + q_{i,c})c + r_{i,c},
\end{aligned}$$

we observe that the smallest message in $A_{t_{j+1}}$ is exactly one larger than the largest message in A_{t_j} , i.e.,

$$\begin{aligned}
&(j + 1)q_{d,c} + q_{i,c} + 1 - (q_{d,c} - 1) \\
&= j q_{d,c} + q_{i,c} + 1 - (q_{d,c} - 1) + q_{d,c} \\
&= (j q_{d,c} + q_{i,c} + 1) + 1.
\end{aligned}$$

Thus, there are no overlapping or omitted messages among the sets of active messages corresponding to $T_n^{(i)}$. Properties 1 and 2 therefore follow.

The total amount of link bandwidth over all time steps in $T_n^{(i)}$, i.e., $|T_n^{(i)}|$, can be computed by summing over the link bandwidth allocated to the n messages, and adding the unused link bandwidth in the smallest time step (which is allocated to nonpositive dummy messages) and in the largest time step (which is allocated to messages larger than n); this produces the required upper bound of Property 3. ■

Proof of Theorem 1: Consider the code constructed in Section III for a given choice of (c, d) . According to Lemma 1, if message size s satisfies the inequality

$$s \leq \sum_{j=1}^{d-z} y_j,$$

then each message $k \in \{1, \dots, n\}$ can be decoded from the data at any $d - z$ time steps in its coding window W_k . Therefore, the code achieves a message size of $\sum_{j=1}^{d-z} y_j$, by allowing all n messages $\{1, \dots, n\}$ to be decoded by their respective deadlines as long as there are z or fewer erasures in each coding window W_k , or equivalently, under any erasure pattern $E \in \mathcal{E}_n^{\text{CW}}$. To demonstrate the asymptotic optimality of the code, we will show that this message size matches the maximum achievable message size s_n^{CW} in the limit, i.e.,

$$\lim_{n \rightarrow \infty} s_n^{\text{CW}} = \sum_{j=1}^{d-z} y_j. \tag{7}$$

To obtain an upper bound for s_n^{CW} , we consider the cut-set bound corresponding to a specific erasure pattern E' from $\mathcal{E}_n^{\text{CW}}$. Let $\{1, \dots, d\}$ be partitioned into two sets $V^{(1)}$ and $V^{(2)}$, where

$$\begin{aligned}
V^{(1)} &\triangleq \{i \in \{1, \dots, d\} : r_{i,c} \leq r_{d,c}\}, \\
V^{(2)} &\triangleq \{i \in \{1, \dots, d\} : r_{i,c} > r_{d,c}\}.
\end{aligned}$$

Let $\mathbf{v} = (v_1, \dots, v_d)$ be defined as $\mathbf{v} \triangleq (\mathbf{v}^{(1)} \mid \mathbf{v}^{(2)})$, where $\mathbf{v}^{(1)}$ is the vector containing the $(q_{d,c} + 1)r_{d,c}$ elements of $V^{(1)}$ sorted in ascending order, and $\mathbf{v}^{(2)}$ is the vector containing the $q_{d,c}(c - r_{d,c})$ elements of $V^{(2)}$ sorted in ascending order. Define the erasure pattern $E' \subseteq T_n$ as follows:

$$E' \triangleq \bigcup_{j=d-z+1}^d T_n^{(v_j)},$$

where $T_n^{(i)}$ is as defined in Lemma 2. The erased time steps in E' have been chosen to coincide with the larger blocks allocated to each message in the constructed code. To show that E' is an admissible erasure pattern, we introduce the following lemma:

Lemma 3. *If $A \subseteq \{1, \dots, d\}$, then*

$$\left| \left(\bigcup_{i \in A} T_n^{(i)} \right) \cap W_k \right| = |A| \quad \forall k \in \{1, \dots, n\}, \tag{8}$$

where $T_n^{(i)}$ is as defined in Lemma 2.

Proof of Lemma 3: Since the code constructed in Section III allocates a positive amount of link bandwidth to each message $k \in \{1, \dots, n\}$ at every time step in its coding window W_k , it follows from Property 2 of Lemma 2 that for each $i \in \{1, \dots, d\}$, we have

$$|T_n^{(i)} \cap W_k| = 1 \quad \forall k \in \{1, \dots, n\}.$$

Equation (8) therefore follows from the fact that $T_n^{(1)}, \dots, T_n^{(d)}$ are disjoint sets. ■

Applying Lemma 3 with $A = \{v_j\}_{j=d-z+1}^d$ produces

$$|E' \cap W_k| = z \quad \forall k \in \{1, \dots, n\},$$

and thus E' is an admissible erasure pattern, i.e., $E' \in \mathcal{E}_n^{\text{CW}}$.

Now, consider a code that achieves the maximum message size s_n^{CW} . Such a code must allow all n messages $\{1, \dots, n\}$ to be decoded under the specific erasure pattern E' . We therefore have the following cut-set bound for s_n^{CW} :

$$n s_n^{\text{CW}} \leq |T_n \setminus E'| \iff s_n^{\text{CW}} \leq \frac{1}{n} |T_n \setminus E'| = \frac{1}{n} \sum_{j=1}^{d-z} |T_n^{(v_j)}|.$$

Applying the upper bounds in Property 3 of Lemma 2, and writing the resulting expression in terms of y_j produces

$$s_n^{\text{CW}} \leq \frac{1}{n} \sum_{j=1}^{d-z} |T_n^{(v_j)}| \leq \frac{1}{n} \sum_{j=1}^{d-z} (n y_j + 2).$$

Since a message size of $\sum_{j=1}^{d-z} y_j$ is known to be achievable (by the constructed code), we have the following upper and lower bounds for s_n^{CW} :

$$\sum_{j=1}^{d-z} y_j \leq s_n^{\text{CW}} \leq \frac{1}{n} \sum_{j=1}^{d-z} (n y_j + 2).$$

These turn out to be matching bounds in the limit as $n \rightarrow \infty$:

$$\sum_{j=1}^{d-z} y_j \leq \lim_{n \rightarrow \infty} s_n^{\text{CW}} \leq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^{d-z} (n y_j + 2) = \sum_{j=1}^{d-z} y_j.$$

We therefore have (7) as required. \blacksquare

Proof of Theorem 2: Consider the code constructed in Section III for a given choice of (c, d) . According to Lemma 1, if message size s satisfies the inequality

$$s \leq \sum_{j=1}^{d-z} y_j,$$

then each message $k \in \{1, \dots, n\}$ can be decoded from the data at any $d-z$ time steps in its coding window W_k . Therefore, the code achieves a message size of $\sum_{j=1}^{d-z} y_j$, by allowing all n messages $\{1, \dots, n\}$ to be decoded by their respective deadlines as long as there are z or fewer erasures in each coding window W_k , which is indeed the case when there are z or fewer erasures in each sliding window L_t , or equivalently, under any erasure pattern $E \in \mathcal{E}_n^{\text{SW}}$. To demonstrate the asymptotic optimality of the code, we will show that this message size matches the maximum achievable message size s_n^{SW} in the limit, i.e.,

$$\lim_{n \rightarrow \infty} s_n^{\text{SW}} = \sum_{j=1}^{d-z} y_j. \quad (9)$$

We consider two cases separately, depending on whether d is a multiple of c :

Case 1: Suppose that d is a multiple of c . In this case, the message size achieved by the constructed code simplifies to

$$\sum_{j=1}^{d-z} y_j = \frac{d-z}{q_{d,c}+1} = \frac{d-z}{d} c.$$

To obtain an upper bound for s_n^{SW} , we consider the cut-set bound corresponding to a specific periodic erasure pattern $E' \subseteq T_n$ given by

$$E' \triangleq \{j d + i \in T_n : j \in \mathbb{Z}_0^+, i \in \{1, \dots, z\}\}.$$

Since E' comprises alternating intervals of z erased time steps and $d-z$ unerased time steps, we have exactly z erasures in each sliding window L_t ; therefore, E' is an admissible erasure pattern, i.e., $E' \in \mathcal{E}_n^{\text{SW}}$.

Now, consider a code that achieves the maximum message size s_n^{SW} . Such a code must allow all n messages $\{1, \dots, n\}$ to be decoded under the specific erasure pattern E' . We therefore have the following cut-set bound for s_n^{SW} :

$$n s_n^{\text{SW}} \leq |T_n \setminus E'| = \left\lfloor \frac{(n-1)c + d}{d} \right\rfloor (d-z) + \max(r' - z, 0),$$

where

$$r' \triangleq (n-1)c + d - \left\lfloor \frac{(n-1)c + d}{d} \right\rfloor d.$$

Further simplification produces

$$s_n^{\text{SW}} \leq \frac{1}{n} \frac{(n-1)c + 2d}{d} (d-z) = \frac{d-z}{d} \left(c + \frac{2d-c}{n} \right).$$

Since a message size of $\frac{d-z}{d}c$ is known to be achievable (by the constructed code), we have the following upper and lower bounds for s_n^{SW} :

$$\frac{d-z}{d}c \leq s_n^{\text{SW}} \leq \frac{d-z}{d} \left(c + \frac{2d-c}{n} \right).$$

These turn out to be matching bounds in the limit as $n \rightarrow \infty$:

$$\frac{d-z}{d}c \leq \lim_{n \rightarrow \infty} s_n^{\text{SW}} \leq \lim_{n \rightarrow \infty} \frac{d-z}{d} \left(c + \frac{2d-c}{n} \right) = \frac{d-z}{d}c.$$

We therefore have (9) as required.

Case 2: Suppose that d is not a multiple of c . Consider a specific base erasure pattern $E' \subseteq T_n$ given by

$$E' \triangleq \bigcup_{j=d-z+1}^d T_n^{(v_j)},$$

where $T_n^{(i)}$ is as defined in Lemma 2, and $\mathbf{v} = (v_1, \dots, v_d)$ is as defined in the proof of Theorem 1. The erased time steps in E' have been chosen to coincide with the larger blocks allocated to each message in the constructed code. From E' , we derive the erasure patterns E'_1, \dots, E'_n given by

$$E'_k \triangleq E' \cap W_k = \bigcup_{j=d-z+1}^d (T_n^{(v_j)} \cap W_k).$$

Applying Lemma 3 with $A = \{v_j\}_{j=d-z+1}^d$ produces

$$|E'_k| = |E' \cap W_k| = z \quad \forall k \in \{1, \dots, n\},$$

which implies that

$$|E'_k \cap L_t| \leq z \quad \forall t \in \{1, \dots, (n-1)c+1\}$$

for each $k \in \{1, \dots, n\}$. Thus, E'_k is an admissible erasure pattern, i.e., $E'_k \in \mathcal{E}_n^{\text{SW}}$, for each $k \in \{1, \dots, n\}$.

To obtain an upper bound for s_n^{SW} , we introduce the following lemma:

Lemma 4. Suppose that a code achieves a message size of s under a given set of erasure patterns \mathcal{E} for a given choice of (n, c, d) . Let X_t be a random variable representing the coded data transmitted at time step $t \in T_n$, let M_k be a random variable representing message $k \in \{1, \dots, n\}$, and define $X[A] \triangleq (X_t)_{t \in A}$. If $E \subseteq T_n$ is such that $E \cap W_k$ is an admissible erasure pattern, i.e., $(E \cap W_k) \in \mathcal{E}$, for each $k \in \{1, \dots, n\}$, then for each $k \in \{1, \dots, n\}$,

$$H(X[W_k \setminus E] \mid M_1^k, X_1^{(k-1)c}) \leq |T_k \setminus E| - ks. \quad (10)$$

Proof of Lemma 4: We will prove by induction that inequality (10) holds for any $k \in \{1, \dots, n\}$.

(Base case) Consider the case of $k = 1$. From the definition of mutual information, we have

$$\begin{aligned} I(X[W_1 \setminus E]; M_1) &= H(X[W_1 \setminus E]) - H(X[W_1 \setminus E] \mid M_1) \\ &= H(M_1) - H(M_1 \mid X[W_1 \setminus E]). \end{aligned}$$

Rearranging terms produces

$$\begin{aligned} H(X[W_1 \setminus E] \mid M_1) &= H(X[W_1 \setminus E]) - H(M_1) \\ &\quad + H(M_1 \mid X[W_1 \setminus E]). \end{aligned} \quad (11)$$

Since $T_1 = W_1$ and $H(X_t) \leq 1$ for any t because of the unit link bandwidth, we have

$$H(X[W_1 \setminus E]) = H(X[T_1 \setminus E]) \leq |T_1 \setminus E|. \quad (12)$$

Furthermore, since $E \cap W_1$ is an admissible erasure pattern, message 1 must be decodable from the coded data at time steps $T_1 \setminus (E \cap W_1) = W_1 \setminus (E \cap W_1) = W_1 \setminus E$, and so

$$H(M_1 \mid X[W_1 \setminus E]) = 0. \quad (13)$$

Substituting (12), (13), and $H(M_1) = s$ into (11) yields

$$H(X[W_1 \setminus E] \mid M_1) \leq |T_1 \setminus E| - s,$$

as required.

(Inductive step) Suppose that

$$H(X[W_k \setminus E] \mid M_1^k, X_1^{(k-1)c}) \leq |T_k \setminus E| - ks \quad (14)$$

for some $k \in \{1, \dots, n-1\}$. From the definition of conditional mutual information, we have

$$\begin{aligned} I(X[W_{k+1} \setminus E]; M_{k+1} \mid M_1^k, X_1^{kc}) \\ &= H(X[W_{k+1} \setminus E] \mid M_1^k, X_1^{kc}) \\ &\quad - H(X[W_{k+1} \setminus E] \mid M_1^{k+1}, X_1^{kc}) \end{aligned}$$

$$\begin{aligned} &= H(M_{k+1} \mid M_1^k, X_1^{kc}) \\ &\quad - H(M_{k+1} \mid M_1^k, X[\{1, \dots, kc\} \cup (W_{k+1} \setminus E)]). \end{aligned}$$

Rearranging terms produces

$$\begin{aligned} &H(X[W_{k+1} \setminus E] \mid M_1^{k+1}, X_1^{kc}) \\ &= H(X[W_{k+1} \setminus E] \mid M_1^k, X_1^{kc}) \\ &\quad - H(M_{k+1} \mid M_1^k, X_1^{kc}) \\ &\quad + H(M_{k+1} \mid M_1^k, X[\{1, \dots, kc\} \cup (W_{k+1} \setminus E)]). \end{aligned} \quad (15)$$

Since messages are independent and message $k+1$ is created at time step $kc+1$, we have

$$H(M_{k+1} \mid M_1^k, X_1^{kc}) = H(M_{k+1}) = s. \quad (16)$$

Furthermore, since $E \cap W_{k+1}$ is an admissible erasure pattern, message $k+1$ must be decodable from the coded data at time steps $T_{k+1} \setminus (E \cap W_{k+1}) = (T_{k+1} \setminus W_{k+1}) \cup (W_{k+1} \setminus E) = \{1, \dots, kc\} \cup (W_{k+1} \setminus E)$, and so

$$H(M_{k+1} \mid M_1^k, X[\{1, \dots, kc\} \cup (W_{k+1} \setminus E)]) = 0. \quad (17)$$

Substituting (16) and (17) into (15) yields

$$\begin{aligned} &H(X[W_{k+1} \setminus E] \mid M_1^{k+1}, X_1^{kc}) \\ &= H(X[W_{k+1} \setminus E] \mid M_1^k, X_1^{kc}) - s \\ &\stackrel{(a)}{\leq} H(X[(W_k \setminus E) \cup (W_{k+1} \setminus E)] \mid M_1^k, X_1^{kc}) - s \\ &\stackrel{(b)}{\leq} H(X[(W_k \setminus E) \cup (W_{k+1} \setminus E)] \mid M_1^k, X_1^{(k-1)c}) - s \\ &\stackrel{(c)}{\leq} H(X[W_k \setminus E] \mid M_1^k, X_1^{(k-1)c}) \\ &\quad + H(X[(W_{k+1} \setminus E) \setminus (W_k \setminus E)]) - s \\ &\stackrel{(d)}{\leq} |T_k \setminus E| - ks + |(W_{k+1} \setminus E) \setminus (W_k \setminus E)| - s \\ &\stackrel{(e)}{=} |T_{k+1} \setminus E| - (k+1)s, \end{aligned}$$

as required, where

- (a) follows from the addition of random variables $X[W_k \setminus E]$ in the entropy term;
- (b) follows from the removal of conditioned random variables $X_{(k-1)c+1}^{kc}$ in the entropy term;
- (c) follows from the chain rule for joint entropy, and the removal of conditioned random variables $X[W_k \setminus E]$, M_1^k , and $X_1^{(k-1)c}$ in the second entropy term;
- (d) follows from the inductive hypothesis (14), and the fact that $H(X_t) \leq 1$ for any t because of the unit link bandwidth;
- (e) follows from the fact that

$$\begin{aligned} &|T_k \setminus E| + |(W_{k+1} \setminus E) \setminus (W_k \setminus E)| \\ &= |T_k \setminus E| + |(W_{k+1} \setminus W_k) \setminus E| \end{aligned}$$

$$\begin{aligned}
&= |T_k \setminus E| + |(T_{k+1} \setminus T_k) \setminus E| \\
&= |T_{k+1} \setminus E|.
\end{aligned}$$

Applying Lemma 4 with $\mathcal{E} = \mathcal{E}_n^{\text{SW}}$ and $E = E'$ to an optimal code that achieves a message size of s_n^{SW} produces

$$H(X[W_k \setminus E'] | M_1^k, X_1^{(k-1)c}) \leq |T_k \setminus E'| - k s_n^{\text{SW}}$$

for any $k \in \{1, \dots, n\}$. Since the conditional entropy term is nonnegative, it follows that for the choice of $k = n$, we have

$$|T_n \setminus E'| - n s_n^{\text{SW}} \geq 0 \iff s_n^{\text{SW}} \leq \frac{1}{n} |T_n \setminus E'| = \frac{1}{n} \sum_{j=1}^{d-z} |T_n^{(v_j)}|.$$

Applying the upper bounds in Property 3 of Lemma 2, and writing the resulting expression in terms of y_j produces

$$s_n^{\text{SW}} \leq \frac{1}{n} \sum_{j=1}^{d-z} |T_n^{(v_j)}| \leq \frac{1}{n} \sum_{j=1}^{d-z} (n y_j + 2).$$

Since a message size of $\sum_{j=1}^{d-z} y_j$ is known to be achievable (by the constructed code), we have the following upper and lower bounds for s_n^{SW} :

$$\sum_{j=1}^{d-z} y_j \leq s_n^{\text{SW}} \leq \frac{1}{n} \sum_{j=1}^{d-z} (n y_j + 2).$$

These turn out to be matching bounds in the limit as $n \rightarrow \infty$:

$$\sum_{j=1}^{d-z} y_j \leq \lim_{n \rightarrow \infty} s_n^{\text{SW}} \leq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^{d-z} (n y_j + 2) = \sum_{j=1}^{d-z} y_j.$$

We therefore have (9) as required. \blacksquare

Proof of Theorem 3: Observe that under each erasure pattern $E \in \mathcal{E}_n^{\text{B}}$, the coding window W_k for each message $k \in \{1, \dots, n\}$ contains at most z erasures: if W_k intersects with zero erasure bursts, then it contains zero erasures; if W_k intersects with exactly one erasure burst, then it contains at most z erasures, i.e., the maximum length of a burst; if W_k intersects with two or more erasure bursts, then it contains a gap of at least $d - z$ unerased time steps between consecutive bursts, and therefore contains at most z erasures.

Consider the code constructed in Section III for a given choice of (c, d) . According to Lemma 1, if message size s satisfies the inequality

$$s \leq \sum_{j=1}^{d-z} y_j,$$

then each message $k \in \{1, \dots, n\}$ can be decoded from the data at any $d - z$ time steps in its coding window W_k . Therefore, the code achieves a message size of $\sum_{j=1}^{d-z} y_j$, by allowing all n messages $\{1, \dots, n\}$ to be decoded by their respective deadlines as long as there are z or fewer erasures in each coding window W_k , which is indeed the case under any erasure pattern $E \in \mathcal{E}_n^{\text{B}}$. To demonstrate the asymptotic optimality of the code, we will show that this message size

matches the maximum achievable message size s_n^{B} in the limit, i.e.,

$$\lim_{n \rightarrow \infty} s_n^{\text{B}} = \sum_{j=1}^{d-z} y_j, \quad (18)$$

for the following three cases:

Case 1: Suppose that d is a multiple of c . In this case, the message size achieved by the constructed code simplifies to

$$\sum_{j=1}^{d-z} y_j = \frac{d - z}{q_{d,c} + 1} = \frac{d - z}{d} c.$$

To obtain an upper bound for s_n^{B} , we consider the cut-set bound corresponding to a specific periodic erasure pattern $E' \subseteq T_n$ given by

$$E' \triangleq \{j d + i \in T_n : j \in \mathbb{Z}_0^+, i \in \{1, \dots, z\}\}.$$

Since E' comprises alternating intervals of z erased time steps and $d - z$ unerased time steps, it is an admissible erasure pattern, i.e., $E' \in \mathcal{E}_n^{\text{B}}$.

The rest of the proof leading to the obtainment of (18) is the same as that of Case 1 in the proof of Theorem 2, with s_n^{SW} replaced by s_n^{B} .

Case 2: Suppose that d is not a multiple of c , and $z \leq c - r_{d,c}$. In this case, the message size achieved by the constructed code simplifies to

$$\sum_{j=1}^{d-z} y_j = c - \sum_{j=d-z+1}^d y_j = c - \frac{z}{q_{d,c}}.$$

Consider a specific *base* erasure pattern $E' \subseteq T_n$ given by

$$E' \triangleq \bigcup_{j=d-z+1}^d T_n^{(v_j)},$$

where $T_n^{(i)}$ is as defined in Lemma 2, and $\mathbf{v} = (v_1, \dots, v_d)$ is as defined in the proof of Theorem 1. The erased time steps in E' have been chosen to coincide with the larger blocks allocated to each message in the constructed code. In this case, E' simplifies to

$$\begin{aligned}
E' &= \bigcup_{r_{i,c}=c-z+1}^c T_n^{((q_{d,c}-1)c+r_{i,c})} \\
&= \left\{ ((j+1)q_{d,c} - 1)c + r_{i,c} \in T_n : \right. \\
&\quad \left. j \in \mathbb{Z}_0^+, r_{i,c} \in \{c - z + 1, \dots, c\} \right\},
\end{aligned}$$

which follows from the definition of $T_n^{(i)}$ and the fact that $r_{i,c} > r_{d,c}$ when $r_{i,c} \in \{c - z + 1, \dots, c\}$. Observe that E' comprises alternating intervals of z erased time steps and $q_{d,c}c - z$ unerased time steps, with each interval of erased time steps corresponding to a specific choice of $j \in \mathbb{Z}_0^+$. Since each erased time step $t \in E'$ can be expressed as

$$t = \underbrace{((j+1)q_{d,c} - 1)c}_{q_{t,c}} + \underbrace{r_{i,c}}_{r_{t,c}},$$

it follows from Section III that the set of active messages A_t at time step t is given by

$$A_t = \left\{ \underbrace{(j+1)q_{d,c} - (q_{d,c} - 1)}_{q_{t,c}+1}, \dots, \underbrace{(j+1)q_{d,c}}_{q_{t,c}+1} \right\}.$$

Therefore, the set of active messages A_t is the same at every time step t in a given interval of z erased time steps (corresponding to a specific j).

From E' , we derive the erasure patterns E'_1, \dots, E'_n given by

$$E'_k \triangleq E' \cap W_k = \bigcup_{j=d-z+1}^d \left(T_n^{(v_j)} \cap W_k \right).$$

Applying Lemma 3 with $A = \{v_j\}_{j=d-z+1}^d$ produces

$$|E'_k| = |E' \cap W_k| = z \quad \forall k \in \{1, \dots, n\}.$$

Let $t' \in E'_k$ be one of the z erased time steps in W_k under erasure pattern E'_k . As previously established, t' belongs to an interval of z erased time steps in E' that have the same set of active messages $A_{t'}$ (which contains message k). It follows that this interval of z erased time steps is also in E'_k , and must therefore constitute E'_k itself. Thus, E'_k is an admissible erasure pattern, i.e., $E'_k \in \mathcal{E}_n^B$, for each $k \in \{1, \dots, n\}$, because it comprises a single erasure burst of z time steps.

Applying Lemma 4 with $\mathcal{E} = \mathcal{E}_n^B$ and $E = E'$ to an optimal code that achieves a message size of s_n^B produces

$$H\left(X[W_k \setminus E'] \mid M_1^k, X_1^{(k-1)c}\right) \leq |T_k \setminus E'| - k s_n^B$$

for any $k \in \{1, \dots, n\}$. Since the conditional entropy term is nonnegative, it follows that for the choice of $k = n$, we have

$$|T_n \setminus E'| - n s_n^B \geq 0 \iff s_n^B \leq \frac{1}{n} |T_n \setminus E'| = \frac{1}{n} \sum_{j=1}^{d-z} |T_n^{(v_j)}|.$$

The rest of the proof leading to the obtainment of (18) is the same as that of Case 2 in the proof of Theorem 2, with s_n^B replaced by s_n^B .

Case 3: Suppose that d is not a multiple of c , and $z \geq d - r_{d,c} = q_{d,c}c$. In this case, the message size achieved by the constructed code simplifies to

$$\sum_{j=1}^{d-z} y_j = \frac{d-z}{q_{d,c}+1}.$$

Consider a specific *base* erasure pattern $E' \subseteq T_n$ given by

$$E' \triangleq \bigcup_{j=d-z+1}^d T_n^{(v_j)},$$

where $T_n^{(i)}$ is as defined in Lemma 2, and $\mathbf{v} = (v_1, \dots, v_d)$ is as defined in the proof of Theorem 1. The erased time steps in E' have been chosen to coincide with the larger blocks allocated to each message in the constructed code. In this case, E' simplifies to

$$E' = T_n \setminus \left(\bigcup_{r_{i,c}=1}^{d-z} T_n^{(r_{i,c})} \right)$$

$$= T_n \setminus \left\{ (j(q_{d,c}+1))c + r_{i,c} \in T_n : j \in \mathbb{Z}_0^+, r_{i,c} \in \{1, \dots, d-z\} \right\},$$

which follows from the definition of $T_n^{(i)}$ and the fact that $r_{i,c} \leq r_{d,c}$ when $r_{i,c} \in \{1, \dots, d-z\}$. Observe that E' comprises alternating intervals of $d-z$ unerased time steps and $(q_{d,c}+1)c - (d-z) = c - r_{d,c} + z$ erased time steps, with each interval of unerased time steps corresponding to a specific choice of $j \in \mathbb{Z}_0^+$. Since each unerased time step $t \in T_n \setminus E'$ can be expressed as

$$t = \underbrace{(j(q_{d,c}+1))c}_{q_{t,c}} + \underbrace{r_{i,c}}_{r_{t,c}},$$

it follows from Section III that the set of active messages A_t at time step t is given by

$$A_t = \left\{ \underbrace{j(q_{d,c}+1)+1 - q_{d,c}}_{q_{t,c}}, \dots, \underbrace{j(q_{d,c}+1)+1}_{q_{t,c}} \right\}.$$

Therefore, the set of active messages A_t is the same at every time step t in a given interval of $d-z$ unerased time steps (corresponding to a specific j).

From E' , we derive the erasure patterns E'_1, \dots, E'_n given by

$$E'_k \triangleq E' \cap W_k = \bigcup_{j=d-z+1}^d \left(T_n^{(v_j)} \cap W_k \right).$$

Applying Lemma 3 with $A = \{v_j\}_{j=d-z+1}^d$ produces

$$|E'_k| = |E' \cap W_k| = z \quad \forall k \in \{1, \dots, n\}.$$

Let $t' \in W_k \setminus E'_k$ be one of the $d-z$ unerased time steps in W_k under erasure pattern E'_k . As previously established, t' belongs to an interval of $d-z$ unerased time steps in $T_n \setminus E'$ that have the same set of active messages $A_{t'}$ (which contains message k). It follows that this interval of $d-z$ unerased time steps is also in $W_k \setminus E'_k$, and must therefore constitute $W_k \setminus E'_k$ itself. Thus, E'_k is an admissible erasure pattern, i.e., $E'_k \in \mathcal{E}_n^B$, for each $k \in \{1, \dots, n\}$, because it comprises either a single erasure burst of z time steps, or two erasure bursts with a combined length of z time steps separated by a gap of $d-z$ unerased time steps.

Applying Lemma 4 with $\mathcal{E} = \mathcal{E}_n^B$ and $E = E'$ to an optimal code that achieves a message size of s_n^B produces

$$H\left(X[W_k \setminus E'] \mid M_1^k, X_1^{(k-1)c}\right) \leq |T_k \setminus E'| - k s_n^B$$

for any $k \in \{1, \dots, n\}$. Since the conditional entropy term is nonnegative, it follows that for the choice of $k = n$, we have

$$|T_n \setminus E'| - n s_n^B \geq 0 \iff s_n^B \leq \frac{1}{n} |T_n \setminus E'| = \frac{1}{n} \sum_{j=1}^{d-z} |T_n^{(v_j)}|.$$

The rest of the proof leading to the obtainment of (18) is the same as that of Case 2 in the proof of Theorem 2, with s_n^B replaced by s_n^B . \blacksquare

REFERENCES

- [1] D. Leong and T. Ho, "Erasure coding for real-time streaming," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, Massachusetts, USA, Jul. 2012.
- [2] E. Martinian and C.-E. W. Sundberg, "Low delay burst erasure correction codes," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2002, pp. 1736–1740.
- [3] E. Martinian and M. Trott, "Delay-optimal burst erasure code construction," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2007, pp. 1006–1010.
- [4] L. J. Schulman, "Coding for interactive communication," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1745–1756, Nov. 1996.
- [5] A. Sahai, "Anytime information theory," Ph.D. dissertation, Massachusetts Institute of Technology, 2001.
- [6] R. T. Sukhvasi, "Distributed control and computing: Optimal estimation, error correcting codes, and interactive protocols," Ph.D. dissertation, California Institute of Technology, 2012.
- [7] Ö. F. Tekin, S. Vyetrenko, T. Ho, and H. Yao, "Erasure correction for nested receivers," in *Proc. Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2011, pp. 1454–1461.